On the determination of the leptonic CP phase

Jessica Elevant[*], **Thomas Schwetz**[†],

Oskar Klein Centre for Cosmoparticle Physics,
Department of Physics, Stockholm University, SE-10691 Stockholm, Sweden

Abstract

The combination of data from long-baseline and reactor oscillation experiments leads to a preference of the leptonic CP phase $\delta_{\rm CP}$ in the range between π and 2π. We study the statistical significance of this hint by performing a Monte Carlo simulation of the relevant data. We find that the distribution of the standard test statistic used to derive confidence intervals for $\delta_{\rm CP}$ is highly non-Gaussian and depends on the unknown true values of θ_{23} and the neutrino mass ordering. Values of $\delta_{\rm CP}$ around $\pi/2$ are disfavored at between 2σ and 3σ, depending on the unknown true values of θ_{23} and the mass ordering. Typically the standard χ^2 approximation leads to over-coverage of the confidence intervals for $\delta_{\rm CP}$. For the 2-dimensional confidence region in the $(\delta_{\rm CP}, \theta_{23})$ plane the usual χ^2 approximation is better justified. The 2-dimensional region does not include the value $\delta_{\rm CP} = \pi/2$ up to the 86.3% (89.2%) CL assuming a true normal (inverted) mass ordering. Furthermore, we study the sensitivity to $\delta_{\rm CP}$ and θ_{23} of an increased exposure of the T2K experiment, roughly a factor 12 larger than the current exposure and including also anti-neutrino data. Also in this case deviations from Gaussianity may be significant, especially if the mass ordering is unknown.

---

[*]Email: jessica.elevant AT fysik.su.se

[†]Now at: Institut für Kernphysik, Karlsruhe Institute of Technology (KIT), 76021 Karlsruhe, Germany. Email: schwetz AT kit.edu

Contents

1 Introduction

Thanks to a beautiful series of neutrino oscillation experiments [1–15] we have now a clear picture of the mixing pattern in the lepton sector. Neutrino oscillations depend on two neutrino mass-squared differences $\Delta m_{21}^2, \Delta m_{32}^2$ (with $\Delta m_{ij}^2 \equiv m_i^2 - m_j^2$), three mixing angles, $\theta_{12}, \theta_{23}, \theta_{13}$, and one complex phase δ_{CP}, where we adopt the standard parameterization for the leptonic mixing matrix [16]. Out of those six parameters the three mixing angels and the two mass-squared differences are well determined by global data [17–19], up to the sign of Δm_{32}^2 which parametrizes two possible orderings of the neutrino mass states, normal ordering (NO) versus inverted ordering (IO).

One of the ultimate goals of neutrino oscillation physics is to determine the complex phase δ_{CP}. Values of δ_{CP} different from zero and π imply CP violation in the lepton sector [20–22], see refs. [23, 24] for reviews. Determining δ_{CP} and possibly establishing CP violation with reasonable precision is a formidable task which most likely will require high-intensity neutrino beams beyond the current generation of experiments. Nevertheless, already with current experiments, some first hints on a preferred range of δ_{CP} may be obtained at a modest confidence level, see for instance [25–29] for estimates. Indeed, currently available global data seem to indicate a slight preference for the range $\pi < \delta_{CP} < 2\pi$ compared to $0 < \delta_{CP} < \pi$ [17–19]. This hint emerges mostly from the combination of the $\nu_\mu \to \nu_e$ observation from the long-baseline experiments T2K [13] and MINOS [6] with the determination of the mixing angle θ_{13} by reactor experiments DayaBay [12], RENO [10], and DoubleChooz [9].

Assessing the significance of this hint is a non-trivial task. Standard statistical tools usually employed in global fits of neutrino oscillation data are likely to fail for the determination of δ_{CP}. The reason is that several conditions for Wilks theorem [30] are violated in this case: statistics is low, the sensitivity of the data to the parameter is rather poor, and predictions depend non-linearly on the parameter (via trigonometric functions). In the context of present data those issues have been commented on in ref. [17], based on preliminary

simulations of the statistical properties of the used $\Delta\chi^2$ statistics. Similar considerations for future experiments can be found in refs. [31, 32].

In this work we extend the results of ref. [17] and study in detail the distribution of the relevant test statistics by generating large samples of pseudo data and constructing confidence intervals or regions with the correct coverage following the Feldman-Cousins prescription [33]. We study the behavior as a function of the unknown true values of δ_{CP}, θ_{23}, and the neutrino mass ordering. In refs. [31,32] the sensitivity to CP violation has been studied, whereas in this work we concentrate on the related but different problem of constructing confidence intervals for δ_{CP}. In addition to analyzing present data, we also investigate the behavior of the test statistics assuming an increased exposure of the T2K experiment, to be expected in the timescale of several years. We attempt to provide an explanation of our numerical results by considering the non-linear structure of the relevant oscillation probabilities including parameter degeneracies.

The outline of the paper is as follows. In section 2 we briefly describe the data from the T2K and MINOS experiments, introduce the relevant test statistics, and discuss the statistical analysis based on the Monte Carlo simulations. In section 3 we consider the relevant oscillation probabilities and provide a discussion about why deviations from Gaussianity can be expected. Our results analyzing present data are presented in section 4. We discuss the distributions of the 1-dimensional $\Delta\chi^2$ statistics for δ_{CP} and θ_{23}, finding large non-Gaussian behavior, especially for δ_{CP}. Then we construct the 2-dimensional regions in the $(\delta_{CP}, \theta_{23})$ plane and find them to be much closer to the Gaussian approximation. In section 5 we investigate how this situation will change, once more data become available. We study the sensitivity of T2K by increasing the exposure by roughly a factor 12 compared to the present one including also anti-neutrino data. We find even in that situation deviations from the Gaussian approximation remain significant in certain regions of the parameter space. We summarize and conclude in section 6. In the appendix we show the impact of first data on anti-neutrinos from T2K [34], which appeared after the completion of this work.

2 Description of data and statistical analysis

In this work we use the data from the long-baseline experiments T2K and MINOS, both from the appearance and disappearance channels, including also a small anti-neutrino data sample from MINOS, see Tab. 1 for details and references.[1] Our code departs from the re-analysis of the data developed in the context of the NuFit collaboration and used in ref. [17]. For each of the six data samples shown in Tab. 1 we perform a spectral fit, where the numbers of spectral bins are given in the table. Our predictions of the event spectra $T_i^r(\Theta)$ have been calibrated in order to reproduce the expected spectra provided by the collaborations. Here r runs over the six data samples, i labels the energy bins, and Θ collectively denotes the oscillation parameters. Each data set is described by a χ^2 statistics appropriate for Poisson

---

[1]Preliminarry results from a T2K anti-neutrino run have been released recently at EPS HEP 2015 [34], consisting of 3 events in the appearance channel. We comment on the impact of these data in the appendix.

Experiment	Channel	Exposure (p.o.t.)	Ref.	Data points	Events
T2K	$\nu_\mu \to \nu_\mu$	6.57×10^{20}	[14]	16	120
T2K	$\nu_\mu \to \nu_e$	6.57×10^{20}	[13]	5	28
MINOS	$\nu_\mu \to \nu_\mu$	10.71×10^{20}	[7]	39	2782
MINOS	$\overline{\nu}_\mu \to \overline{\nu}_\mu$	3.36×10^{20}	[7]	14	222
MINOS	$\nu_\mu \to \nu_e$	10.6×10^{20}	[6]	5	88
MINOS	$\overline{\nu}_\mu \to \overline{\nu}_e$	3.3×10^{20}	[6]	5	9

Table 1: Summary of used data. The last two columns give the number of bins used to fit the energy spectrum, and the total number of observed events, respectively. Recent T2K data on the $\overline{\nu}_\mu \to \overline{\nu}_e$ channel [34] are not used in the main text, but we show some results including them in the appendix.

distributed data:

$$\chi_r^2(\Theta, a_r) = 2 \sum_i \left[a_r\, T_i^r(\Theta) - O_i^r + O_i^r \log \frac{O_i^r}{a_r\, T_i^r(\Theta)} \right], \tag{1}$$

$$\chi_r^2(\Theta) = \min_{a_r} \left[\chi_r^2(\Theta, a_r) + \left(\frac{1 - a_r}{\sigma_{\text{sys}}^r} \right)^2 \right], \tag{2}$$

where O_i^r is the observed number of events, and σ_{sys}^r is the systematic over-all normalization error included via the pull parameters a_r. Since those data are largely statistics dominated, this simple treatment of systematic errors suffice. For each experiment our analysis is validated by checking that when analyzing the data in the same way we can reproduce the confidence regions in parameter space obtained by the experimental collaborations with good accuracy. When combining the data samples given in Tab. 1 we simply add the χ^2 functions,

$$\chi^2(\Theta) = \sum_r \chi_r^2(\Theta), \tag{3}$$

ignoring possible correlated systematic errors between the data sets. The results from the reactor experiments [9, 10, 12] are taken into account implicitly by fixing $\theta_{13} = 8.5°$.

Below we are going to focus on the parameters δ_{CP}, θ_{23}, and Δm_{32}^2, which currently have the largest uncertainties, including the sign of Δm_{32}^2. The other oscillation parameters are fixed to $\theta_{12} = 33.5°$, $\theta_{13} = 8.5°$, $\Delta m_{21}^2 = 7.5 \times 10^{-5}\,\text{eV}^2$. Those parameters are known within better than 15% at 3σ [2] and we expect that fixing those parameters has only a small impact on our results. Hence, in the notation used above we have $\Theta = \{\delta_{\text{CP}}, \theta_{23}, \Delta m_{32}^2\}$.

If we are interested in confidence regions of one of the parameters $\Theta = \{\delta_{\text{CP}}, \theta_{23}, \Delta m_{32}^2\}$, irrespective of the others, we consider the following test statistic. Taking for example δ_{CP}, we define

$$\Delta\chi^2(\delta_{\text{CP}}) = \min_{\theta_{23}, \Delta m_{32}^2} \chi^2(\Theta) - \chi_{\min}^2, \tag{4}$$

where $\chi_{\min}^2$ is the global minimum of the χ^2 with respect to all parameters Θ. Note that when minimizing over Δm_{32}^2, we always take into account both signs, i.e. we minimize also over the two mass orderings. Similar definitions apply for the other 1-dimensional cases,

---

[2] Here we define the precision by $2(x^{\text{up}} - x^{\text{low}})/(x^{\text{up}} + x^{\text{low}})$, where x^{up} and x^{low} are the upper and lower ends of the 3σ intervals [17], respectively.

$\Delta\chi^2(\theta_{23})$ and $\Delta\chi^2(\Delta m_{32}^2)$. The 2-dimensional confidence regions are based on an analogous definition, e.g.,

$$\Delta\chi^2(\delta_{\mathrm{CP}}, \theta_{23}) = \min_{\Delta m_{32}^2} \chi^2(\Theta) - \chi^2_{\min}. \tag{5}$$

This procedure is equivalent to the profile-likelihood method to treat nuisance parameters.

Wilks theorem [30] implies that under certain conditions, the test statistics from eqs. (4) and (5) are distributed according to the χ^2-distribution with 1 and 2 degrees of freedom (dof), respectively. This is the basis of the standard method to derive confidence regions for the parameters, using the condition $\Delta\chi^2 \leq t_{\chi^2}(\mathrm{CL}, \mathrm{dof})$. We refer to t_{χ^2} as "cut levels", and their values can be obtained by integrating the corresponding χ^2 distribution. For instance, the 1-dimensional intervals at $1\sigma, 2\sigma, 3\sigma$ are derived by $\Delta\chi^2 \leq 1, 4, 9$, respectively.[3] We will refer to this situation as the "Gaussian limit" or the "χ^2 limit" in the following.

Wilks theorem applies if the theoretical predictions $T_i(\Theta)$ span a linear space when Θ is varied. For instance, this is the case if $T_i(\Theta)$ can be expanded to linear order: $T_i(\Theta) \approx A_i + B_i\Theta$. This is trivially fulfilled for a linear model, where this relation is exact. For non-linear models $T_i(\Theta)$, the linear approximation will hold in the vicinity of the best fit point and will be reliable up to a certain CL, beyond which the non-linear character of the parameter dependence can lead to deviations from the Gaussian limit. For "powerful" data, which constrain the parameter efficiently, the linear approximation will hold up to a high CL, whereas for "weak" data with poor sensitivity to the parameter it will break down already at low CL. Deviations from Gaussianity are expected for example close to a physical boundary of a parameter, or when certain values of the predictions $T_i(\Theta)$ cannot be reached due to the parameter dependence of the model, for instance via trigonometric functions, as we are going to see below.

In order to study deviations from the Gaussian limit we have performed Monte Carlo (MC) simulations and calculated the distributions of the test statistics numerically. For assumed true values of the parameters Θ^{true} we generate artificial pseudo data by assuming a Poisson distribution for the observables with mean $T_i^r(\Theta^{\mathrm{true}})$. Those data are multiplied by a random Gaussian number with mean 1 and standard deviation σ_{sys}^r in order to take into account the systematic normalization uncertainty. In this approach the origin of the systematic uncertainty is related to some auxiliary measurements, determining for instance the fiducial volume of the detector or the beam normalization. The *measured* values of those experiments are used to determine the theoretical predictions. Hence, the normalization of the predictions is subject to statistical fluctuations of the auxiliary measurements. We set the unknown true value of the normalization constant to 1 and generate random realisations of this number with standard deviation σ_{sys}^r, which then enters the "observables" for the MC generated data. This implies that we consider the auxiliary measurements as part of the experiment, also hypothetically to be repeated many times. However, we have checked that our results do not depend on whether we randomize the systematic error or not.

Then the test statistics eqs. (4) or (5) are calculated at the point Θ^{true}, e.g., $\Delta\chi^2_{\mathrm{MC}[\Theta]}(\delta_{\mathrm{CP}})$, where the subscript $\mathrm{MC}[\Theta]$ indicates that the pseudo data from the Monte Carlo generated at the point Θ are used in the χ^2. We perform this calculation 10^4 times for each point in the Θ^{true} space, which provides us the true distribution of the test statistic for each parameter

---

[3]In this paper we use the two-sided Gaussian convention to convert standard deviations into CL, which implies that $1\sigma, 2\sigma, 3\sigma$ correspond to 68.27%, 95.45%, 99.73% CL, respectively.

value. From those histograms we can obtain the true cut levels $t_{\mathrm{MC}}(\mathrm{CL}, \Theta)$ for a given CL α, by demanding that a fraction α of all pseudo experiments fulfills

$$\Delta \chi^2_{\mathrm{MC}[\Theta]}(\delta_{\mathrm{CP}}) \leq t_{\mathrm{MC}}(\mathrm{CL}, \Theta). \tag{6}$$

Then the correct confidence intervals (or regions) for the parameters are obtained by those values of the parameters for which the test statistics of the real data fulfills $\Delta \chi^2(\delta_{\mathrm{CP}}) \leq t_{\mathrm{MC}}(\mathrm{CL}, \Theta)$. Those intervals have the correct coverage by construction and follow the prescription of Feldman and Cousins [33].

Here we used the test statistic for δ_{CP} as an example, but analogous expressions hold for the other 1-dimensional as well as 2-dimensional test statistics. Note that in the left side of eq. (6), $\Delta \chi^2$ is evaluated at δ_{CP} corresponding to the same value as used to generate the pseudo data. However, although the test statistic for given data depends only on δ_{CP}, the MC results do depend also on the other parameters θ_{23} and Δm^2_{32}. Hence, the confidence intervals for δ_{CP} may depend on the unknown true values of the other parameters, an effect we will indeed observe in our numerical studies presented in the next section.

3 Discussion of oscillation probabilities

In the case of interest, we are facing a complicated parameter dependence of the predictions. We review here the relevant oscillation probabilities through which the parameters enter the event rate predictions. Note that we do not use the approximate expression for the numerical work (which is based on numerical calculations of the full three-flavor probabilities including the matter effect) but they serve well for a qualitative understanding.

Let us define

$$\Delta \equiv \frac{|\Delta m^2_{31}| L}{4E}, \quad A \equiv \left| \frac{2EV}{\Delta m^2_{31}} \right|, \tag{7}$$

where L is the baseline, E is the neutrino energy, and V is the effective matter potential. For the ν_μ disappearance channel we have

$$P_{\mathrm{dis}} \approx \sin^2 2\theta_{23} \sin^2 \Delta, \tag{8}$$

where we show only the leading term and neglect corrections due to Δ_{21} as well as θ_{13}. An approximate expression for the $\nu_\mu \to \nu_e$ oscillation probability, valid for a constant matter density is given by [35, 36]:

$$\begin{aligned} P_{\mathrm{app}} \approx \ &4 \sin^2 \theta_{13} \sin^2 \theta_{23} \frac{\sin^2 \Delta(1 - asA)}{(1 - asA)^2} \\ &+ s \frac{\Delta m^2_{21}}{|\Delta m^2_{31}|} \sin 2\theta_{13} \sin 2\theta_{12} \sin 2\theta_{23} \cos(s\Delta + a\delta_{\mathrm{CP}}) \frac{\sin \Delta A}{A} \frac{\sin \Delta(1 - asA)}{1 - asA}. \end{aligned} \tag{9}$$

The signs a and s describe the effects of CP-conjugation and the neutrino mass ordering, respectively, with $a = +1$ for neutrinos and $a = -1$ for anti-neutrinos, and $s = \mathrm{sgn}(\Delta m^2_{31})$. The matter effect enters via the parameter A. Numerically one finds for a matter density of 3 g/cm^3

$$A \simeq 0.094 \left(\frac{E}{\mathrm{GeV}} \right) \left(\frac{|\Delta m^2_{31}|}{2.4 \times 10^{-3} \ \mathrm{eV}^2} \right)^{-1}. \tag{10}$$

Hence, for the T2K experiment, with $E \simeq 0.7$ GeV, the matter effect is of order 6% and we can expand eq. (9) also in A, keeping only terms up to first order in A. To simplify the expression further we assume the first oscillation maximum, $\Delta \approx \pi/2$, which is a good approximation for T2K. Introducing the definitions

$$\sin^2 \theta_{23} = \frac{1}{2} + d, \qquad C \equiv \frac{\Delta m_{21}^2 L}{4E} \sin 2\theta_{13} \sin 2\theta_{12} \sin 2\theta_{23}, \qquad (11)$$

eq. (9) becomes for neutrinos and anti-neutrinos

$$P_\nu \approx 2 \sin^2 \theta_{13}(1 + 2d)(1 + 2sA) - C \sin \delta_{\mathrm{CP}} (1 + sA), \qquad (12)$$
$$P_{\bar\nu} \approx 2 \sin^2 \theta_{13}(1 + 2d)(1 - 2sA) + C \sin \delta_{\mathrm{CP}} (1 - sA). \qquad (13)$$

Note that the magnitude of d is constrained by the $\nu_\mu \to \nu_\mu$ disappearance channel, but we are left with the octant degeneracy for θ_{23}, described by a sign ambiguity of d, with $d < 0$ ($d > 0$) corresponding to the first (second) octant for θ_{23}. We are going to consider values in the range $-0.1 \le d \le 0.1$, within the currently 3σ confidence interval. Hence, the free parameters in the problem are the continuous parameter δ_{CP} and the two signs of d and s, i.e. four discrete sign combinations. Note, however, that especially for current data this is an over-simplification, since the uncertainty on $|d|$ is large. Numerically we have $\sin^2 \theta_{13} \approx 0.022$, $C \approx 0.013$ (with a very weak dependence on d for $|d| \lesssim 0.1$), and $A \approx 0.06$. Hence, all terms in eqs. (12) and (13) are of similar order, and both the octant [37] and the mass ordering [38] degeneracies will lead to changes in the predictions of similar size as δ_{CP}. Recent discussions of the δ_{CP} and θ_{23} interplay can be found in refs. [29, 39, 40].

Given the parameter dependencies from eqs. (8) and (12, 13) we can expect deviations from the Gaussian limit for the following reasons.

1. Present data show only weak sensitivity to δ_{CP}, i.e., the full range $0 \le \delta_{\mathrm{CP}} < 2\pi$ is allowed at relatively low CL. This implies that the strong non-linearity of the trigonometric dependence in eq. (12) comes into play. In particular, because of the sine dependence, only a finite change of $T_i(\delta_{\mathrm{CP}})$ can be achieved by varying δ_{CP}, in contrast to the unbounded variation of a linear model. This means that δ_{CP} effectively provides less than 1 dof, which implies that the true cut levels for a given CL will be lower than the ones corresponding to the Gaussian approximation.

   For future data with decreased statistical errors the fluctuations of the data may become of similar size as the compact region spanned in T_i upon varying δ_{CP}, which means that the coverage of T_i is actually more efficient than for the linear model and δ_{CP} thus provides more than 1 dof. If the exposure is further increased the fluctuations will become even smaller than the compact region in T_i such that the linear expansion becomes valid and we approach 1 dof.

2. Disappearance data depend on $\sin^2 2\theta_{23}$, which leads to the octant degeneracy for θ_{23}. However, due to the $\sin^2 \theta_{23}$ factor in the appearance probability the degeneracy is not complete when $\nu_\mu \to \nu_e$ data are included. The two possible solutions for $\sin^2 \theta_{23}$ corresponding to the two signs of d imply more freedom than a single linear parameter. Hence the presence of the degeneracy leads to an increase of the effective dofs, which implies increased cut levels.

3. For values of θ_{23} close to $\pi/4$ we are facing a physical boundary in the disappearance channel, since $\sin^2 2\theta_{23} \leq 1$, which implies that predictions corresponding to $\sin^2 2\theta_{23} > 1$ cannot be reached by varying θ_{23}. This leads to a decrease of the effective dof, and reduces the cut levels.

4. Similarly, for $\delta_{\rm CP} \simeq \pi/2$ or $3\pi/2$ the $\sin\delta_{\rm CP}$ dependence in the appearance probabilities (12) and (13) imply a physical boundary due to $|\sin\delta_{\rm CP}| \leq 1$. For those values of $\delta_{\rm CP}$ the derivatives of the probabilities with respect to $\delta_{\rm CP}$ vanish. Hence, we expect decreased cut levels for those values, while for $\delta_{\rm CP} \simeq 0$ or π the dependence of the appearance channel resembles approximately a linear model. The behavior around $\delta_{\rm CP} \simeq \pi/2$ or $3\pi/2$ is further complicated by the octant and mass ordering degeneracies, see the discussion related to eq. (14) below.

In our numerical results presented below we will observe all of those effects, where some of them may occur simultaneously, leading to a complicated interplay of effects. Nevertheless, some general features can be understood qualitatively. For instance, considering the $\delta_{\rm CP}$ dependence of eqs. (12) and (13) plus the 4-fold degeneracy related to the signs of d and s, we find that there are minimal and maximal values for the oscillation probabilities (and hence for the event rates) given by the following combinations of the parameters:

$$
\begin{aligned}
P_\nu^{\rm max} : & \quad d > 0 \,(\text{2nd oct.}), & s = +1 \,(\text{NO}), & \quad \delta_{\rm CP} = 3\pi/2 \\
P_\nu^{\rm min} : & \quad d < 0 \,(\text{1st oct.}), & s = -1 \,(\text{IO}), & \quad \delta_{\rm CP} = \pi/2 \\
P_{\bar\nu}^{\rm max} : & \quad d > 0 \,(\text{2nd oct.}), & s = -1 \,(\text{IO}), & \quad \delta_{\rm CP} = \pi/2 \\
P_{\bar\nu}^{\rm min} : & \quad d < 0 \,(\text{1st oct.}), & s = +1 \,(\text{NO}), & \quad \delta_{\rm CP} = 3\pi/2
\end{aligned}
\tag{14}
$$

If the true values of $d, s, \delta_{\rm CP}$ correspond to one of the combinations in eqs. (14) then we are located at a physical boundary for the event rates: there is no point in the parameter space which can provide a larger (or smaller) value of the probability. Statistical fluctuations leading to even larger (or smaller) event rates than predicted for those extreme parameter values cannot be accommodated by adjusting the model. This implies that the effective number of dof of the $\Delta\chi^2$ is reduced, i.e. lower cut levels. A related discussion can also be found in ref. [32].

4 Results for present data

4.1 One-dimensional intervals for $\delta_{\rm CP}$

We start presenting the results of our simulations for the 1-dimensional $\Delta\chi^2$ distributions defined in eq. (4). In figs. 1 and 2 we consider $\Delta\chi^2(\delta_{\rm CP})$ for the CP phase $\delta_{\rm CP}$ and show the cut levels $t_{\rm MC}$ for 1σ, 2σ, 3σ. Fig. 1 uses T2K data only whereas fig. 2 uses all data given in Tab. 1, showing qualitatively similar results to the T2K-only case. We have checked that our T2K results are consistent with the ones shown by the T2K collaboration [15], in cases where comparison is possible.

By comparing the curves for $t_{\rm MC}$ to the χ^2 approximation t_{χ^2} indicated by the dashed curves, we observe significant deviations for the Gaussian limit, with $t_{\rm MC}(2\sigma)$ and $t_{\rm MC}(3\sigma)$ being much lower than the corresponding t_{χ^2} [17]. Furthermore we find large variations of the $\Delta\chi^2(\delta_{\rm CP})$ distribution depending on $\delta_{\rm CP}$ itself, and on the assumed true values for
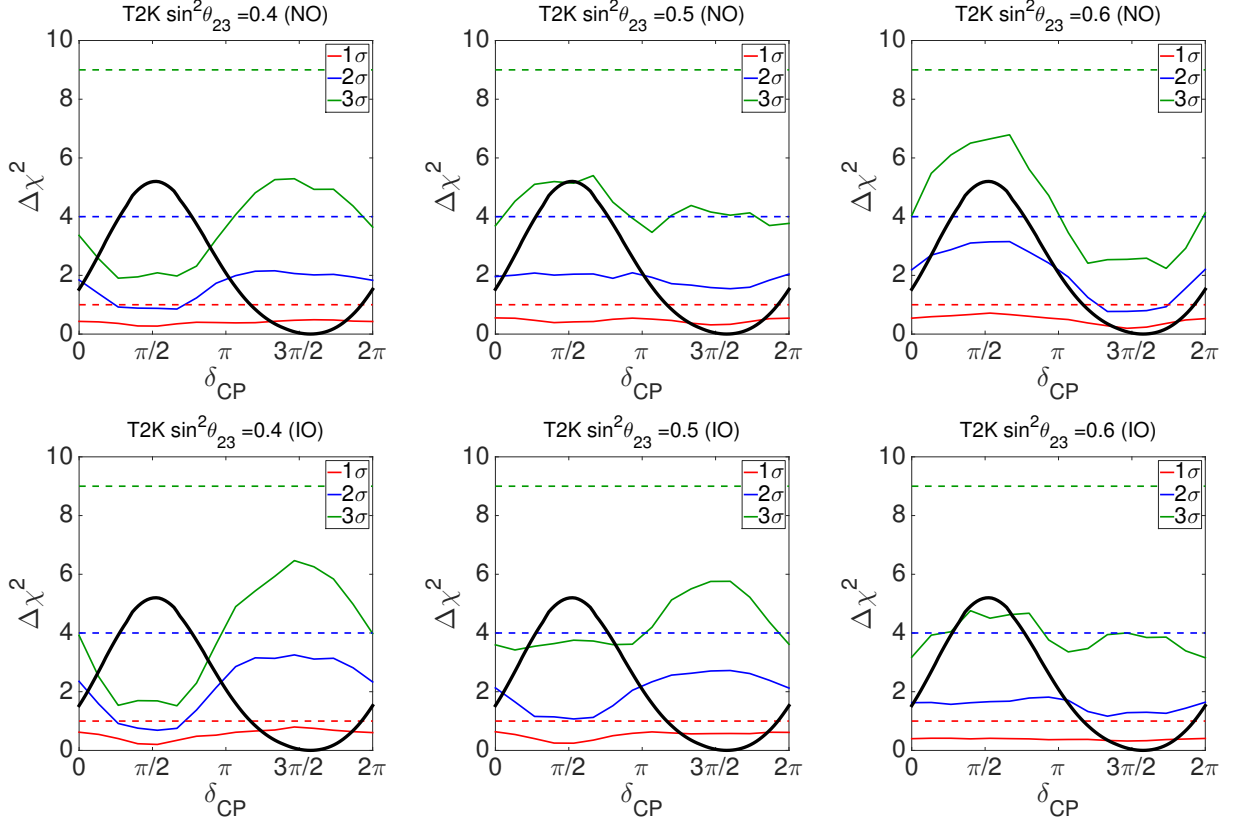
Figure 1: The cut levels $t_{\mathrm{MC}}(\mathrm{CL},\Theta)$ for $\Delta\chi^2(\delta_{\mathrm{CP}})$ from T2K data for 1σ (red), 2σ (blue), 3σ (green). Dashed lines indicate the Gaussian approximation t_{χ^2}. Left, middle, right panels correspond to $\sin^2\theta_{23}^{\mathrm{true}} = 0.4, 0.5, 0.6$, respectively. We take $|\Delta m_{32}^2{}^{\mathrm{true}}| = 2.4 \times 10^{-3}\,\mathrm{eV}^2$ and for the upper (lower) row we have assumed a true normal (inverted) mass ordering. The black solid curve shows $\Delta\chi^2(\delta_{\mathrm{CP}})$ using the observed data (same curve in all panels).

$\sin^2\theta_{23}$ [17] and to a lesser extent also depending on the mass ordering. The various panels in figs. 1 and 2 correspond to different assumptions about $\theta_{23}^{\mathrm{true}}$ and the true mass ordering. Note that those are the "true" values assumed for generating the pseudo data, while when fitting to the data we leave θ_{23} and Δm_{32}^2 free. Comparing figs. 1 and 2, we find that the addition of MINOS data makes some of the "dips" in the cut levels less sever, e.g. the ones for $\sin^2\theta_{23} \simeq 0.4$ and $\delta \simeq \pi/2$ for both orderings.

The behavior of the t_{MC} curves can be understood from the discussion given in section 3. The reduction of t_{MC} compared to the Gaussian limit follows from the poor sensitivity of the data to δ_{CP}, which implies that the full range $0 \le \delta_{\mathrm{CP}} < 2\pi$ becomes accessible. Hence the trigonometric dependence becomes relevant, changing δ_{CP} provides less freedom than a linear parameter, and the effective number of dof becomes reduced.

The appearance of the bumps, for instance for normal ordering and $(\sin^2\theta_{23}^{\mathrm{true}} = 0.4, \delta_{\mathrm{CP}} \simeq 3\pi/2)$ and $(\sin^2\theta_{23}^{\mathrm{true}} = 0.6, \delta_{\mathrm{CP}} \simeq \pi/2)$ can be understood by considering the θ_{23} octant degeneracy. We show in fig. 3 the sensitivity of T2K in the $(\sin^2\theta_{23}, \delta_{\mathrm{CP}})$ plane. We calculate so-called Asimov data, using the theoretical prediction for certain assumed true values without statistical fluctuations as "data". This indicates the expected sensitivity for
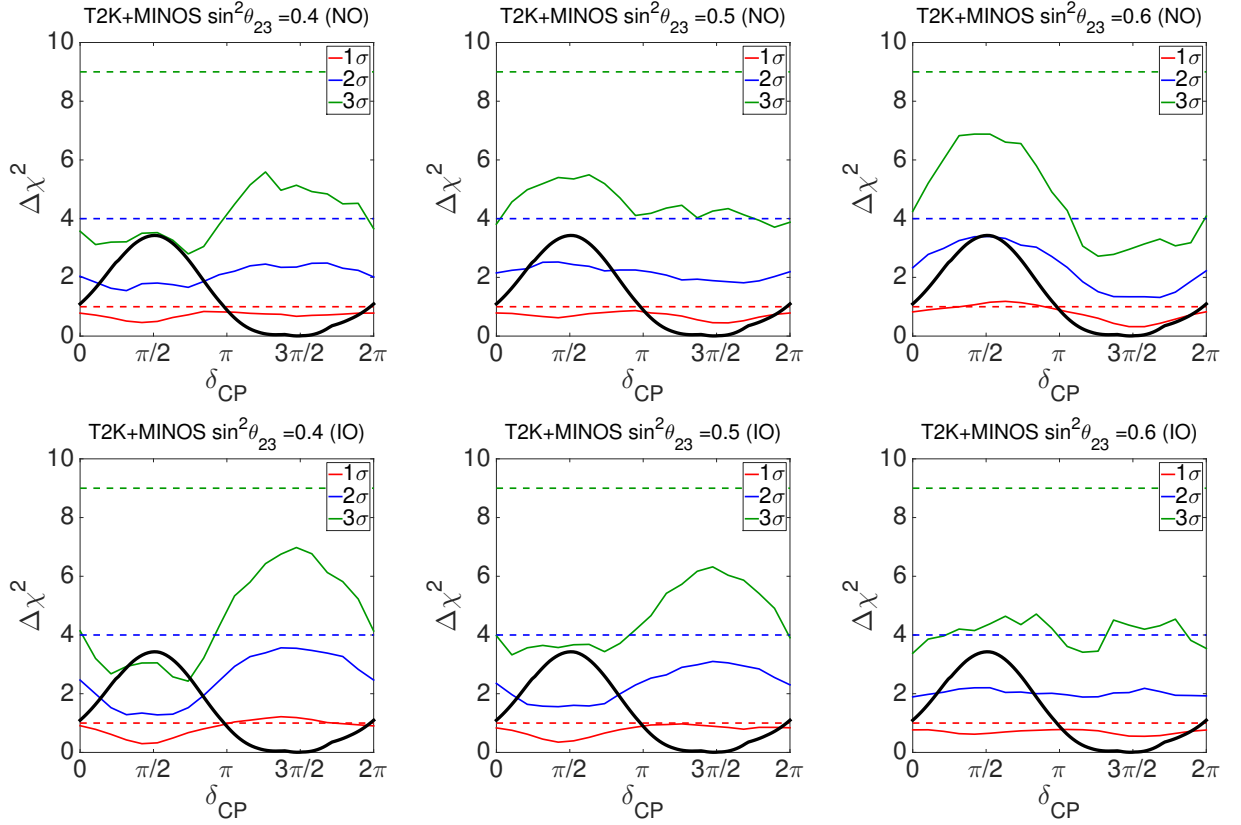
Figure 2: Same as fig. 1 but using combined T2K and MINOS data.

that particular set of true values for an average experiment. Regions in fig. 3 are derived based on the Gaussian approximation for $\Delta\chi^2(\delta_{CP}, \theta_{23})$.

By comparing those results with the upper row of panels in fig. 1 we find a correlation with the ability to resolve the degeneracy: in cases with improved sensitivity to the octant (panels (b) and (h) in fig. 3) the cut levels are low (upper left panel for $\delta_{CP} \simeq \pi/2$ and upper right panel for $\delta_{CP} \simeq 3\pi/2$ in fig. 1), whereas for cases where the degeneracy is strong for all values of δ_{CP} (panels (d) and (f) in fig. 3) the cut levels are high (upper left panel for $\delta_{CP} \simeq 3\pi/2$ and upper right panel for $\delta_{CP} \simeq \pi/2$ in fig. 1). This shows that the presence of the degeneracy increases the effective number of dof for the $\Delta\chi^2$ distribution. We have confirmed a similar correspondence also for the inverted ordering. Note that the overall sensitivity to the octant is very poor even in the cases shown in panels (d) and (f) in fig. 3, corresponding to $\Delta\chi^2 \approx 0.5$ for the wrong octant solution. However, for the distribution of $\Delta\chi^2(\delta_{CP})$ it is relevant that the degeneracy can be resolved for a significant range of δ_{CP} values, while for the the cases corresponding to panels (d) and (f) in fig. 3 the degeneracy is present at below 1σ for all values of δ_{CP}. Those results can also be understood from the discussion related to the maximal and minimal values of the oscillation probability given in the first 2 lines of eq. (14). Fig. 1 shows low cut levels for those combinations of parameters, following from the physical boundary of the event rates.[4]

---

[4]The reduced cut levels for NO, $\sin^2\theta_{23} = 0.4$, $\delta_{CP} \simeq \pi/2$ seen in the left upper panel of fig. 1 do not follow from this argument. While $d < 0$ and $\delta_{CP} \simeq \pi/2$ does correspond to a minimum of the oscillation
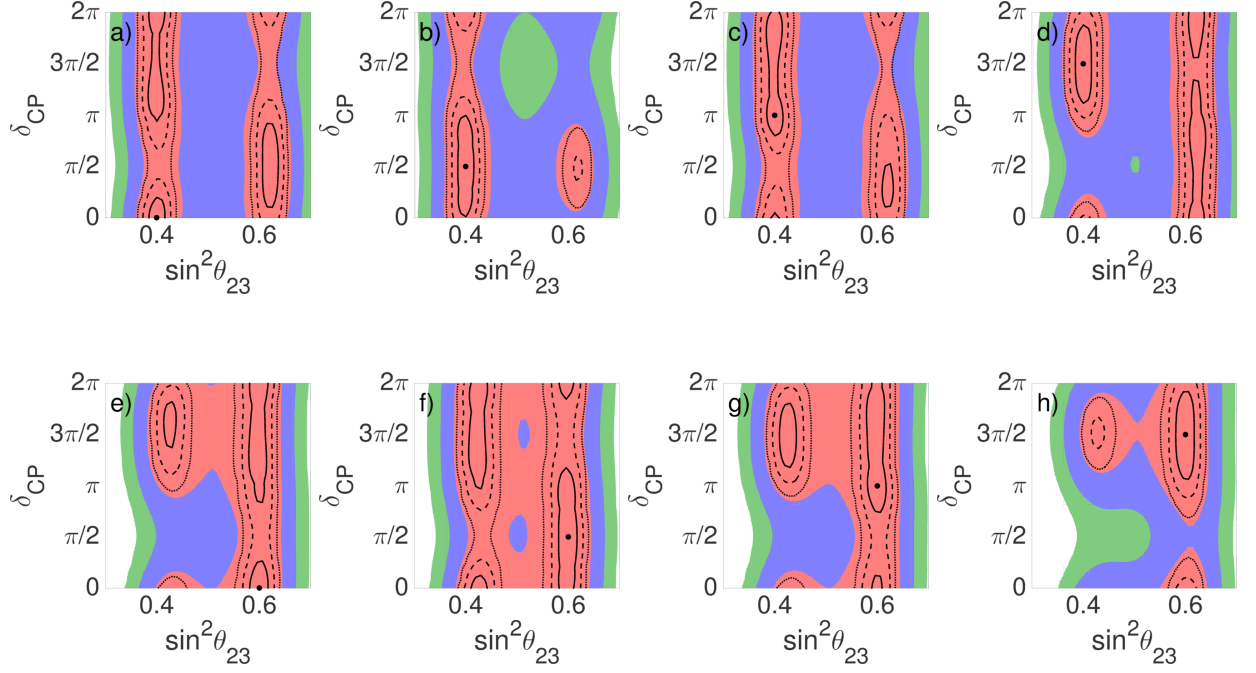
Figure 3: Sensitivity of T2K data in the $(\sin^2\theta_{23}, \delta_{\mathrm{CP}})$ plane based on Asimov data at 1σ, 2σ, 3σ assuming $\sin^2\theta_{23}^{\mathrm{true}} = 0.4$ (upper row) and 0.6 (lower row) for different values of $\delta_{\mathrm{CP}}^{\mathrm{true}}$ as marked by the dots in the plots. We assume $\Delta m_{32}^{2\ \mathrm{true}} = 2.4 \times 10^{-3}\,\mathrm{eV}^2$ (NO). Colored areas correspond to the current exposure, whereas the black contour curves correspond to an exposure of 7.8×10^{21} p.o.t. neutrino data (about 12 times the current exposure). Regions are derived by assuming the Gaussian approximation.

For the cut levels in figs. 1 and 2, we have always minimized with respect to both mass orderings. However, we have checked that assuming that the mass ordering was known (i.e., restricting to the true ordering in the fit) does not change the results qualitatively. This is different from the cut levels for increased exposure discussed in section 5, where we will find a significant effect of restricting the mass ordering.

The exact confidence intervals for δ_{CP} can be obtained by comparing the $\Delta\chi^2(\delta_{\mathrm{CP}})$ from the observed data (shown as black curves in figs. 1 and 2) to the curves for the cut levels. The confidence interval at a given CL is obtained by those values of δ_{CP} for which $\Delta\chi^2(\delta_{\mathrm{CP}}) \leq t_{\mathrm{MC}}(\mathrm{CL})$. We show the confidence intervals for δ_{CP} at 1σ, 2σ, 3σ in table 2 and fig. 4 for different assumptions about the true values of θ_{23} and the mass ordering, and we compare them to the Gaussian approximation. Because of the dependence of t_{MC} on the values of the other parameters the confidence intervals for δ_{CP} depend on those unknown true values. From fig. 4 we see that the 1σ intervals are relatively stable and agree well with the Gaussian approximation. The variations are relatively large for the 2σ interval. There is a large parameter dependence on the CL of rejection of $\delta_{\mathrm{CP}} \simeq \pi/2$. Using combined T2K+MINOS data this rejection ranges from around 3σ for $\sin^2\theta_{23} = 0.4$ (both mass orderings) to only 2σ for $\sin^2\theta_{23} = 0.6$ (NO), see fig. 2. We obtain $\Delta\chi^2 = 3.37$ at $\delta_{\mathrm{CP}} = \pi/2$. Hence, the Gaussian approximation gives a rejection of $\delta_{\mathrm{CP}} = \pi/2$ at 1.8σ. The dependence of the MC results on

---

probability, apparently in that case changing the mass ordering from NO to IH does not provide enough freedom to overcome the physical boundary.

true $\sin^2\theta_{23}$	true MO	1σ	2σ	3σ
0.4	normal	[3.20, 6.00]	[0, 0.42] ∪ [2.65, 2π]	[0, 2π]
0.5	normal	[3.21, 5.94]	[0, 0.70] ∪ [2.48, 2π]	[0, 2π]
0.6	normal	[3.16, 5.90]	[0, 2π]	[0, 2π]
0.4	inverted	[3.11, 6.14]	[0, 0.44] ∪ [2.63, 2π]	[0, 1.08] ∪ [2.35, 2π]
0.5	inverted	[3.15, 6.07]	[0, 0.45] ∪ [2.61, 2π]	[0, 2π]
0.6	inverted	[3.23, 5.92]	[0, 0.55] ∪ [2.59, 2π]	[0, 2π]
Gaussian limit		[3.09, 6.20]	[0, 2π]	[0, 2π]

Table 2: Confidence intervals at 1σ, 2σ, 3σ for the CP phase δ_{CP} from combined T2K and MINOS data for different assumptions about the true values of θ_{23} and the neutrino mass ordering. The last row shows the confidence intervals in the Gaussian approximation.
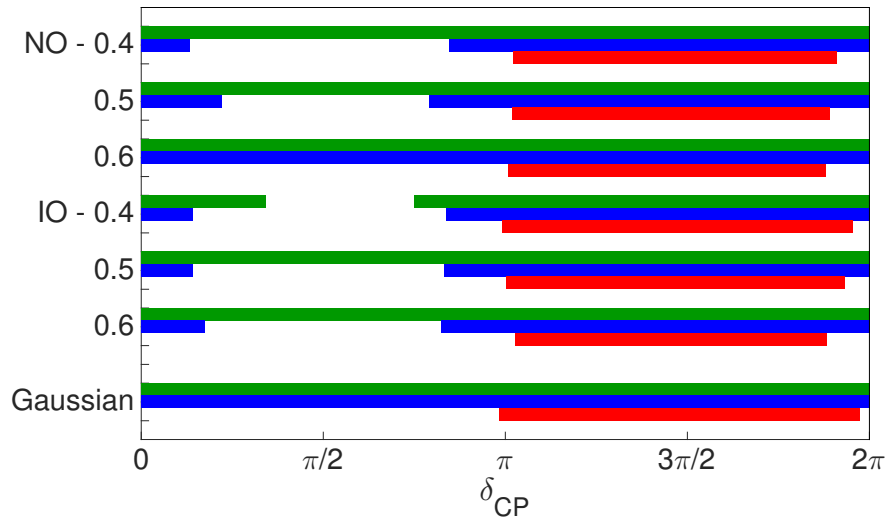

Figure 4: Graphical representation of the confidence intervals for δ_{CP} at 1σ (red), 2σ (blue), 3σ (green), see table 2. The labels on the left side indicate the true value of $\sin^2\theta_{23}$ and the mass ordering. The bottom three bars correspond to the Gaussian approximation.

the true values of other parameters is even more pronounced for T2K data only, see fig. 1. Note that the rejection of $\delta_{CP} \simeq \pi/2$ is stronger for T2K data-only and the significance decreases somewhat when MINOS data are included. The reason for the slight decrease of $\Delta\chi^2(\delta_{CP})$ when MINOS data are added to T2K is that MINOS appearance data prefer a somewhat smaller value of θ_{13} than T2K and hence the combination with the reactor result for θ_{13} becomes somewhat less effective in constraining δ_{CP} when MINOS is included.

Within the frequentist framework there is no way to marginalize or average over the true values of nuisance parameters, since those are considered to be fixed constants of Nature. Hence the dependence of the results for δ_{CP} on the unknown true values of other parameters (especially θ_{23}) introduces an unpleasant ambiguity. One possibility to deal with this situation could be to present for each CL the largest confidence interval for δ_{CP}. Then the CL would be a lower bound on the true coverage of the interval, i.e., such an interval at the α CL will cover the true value with a probability of at least α. A problem with this approach is that one has to maximize the confidence interval of δ_{CP} with respect to the true value of
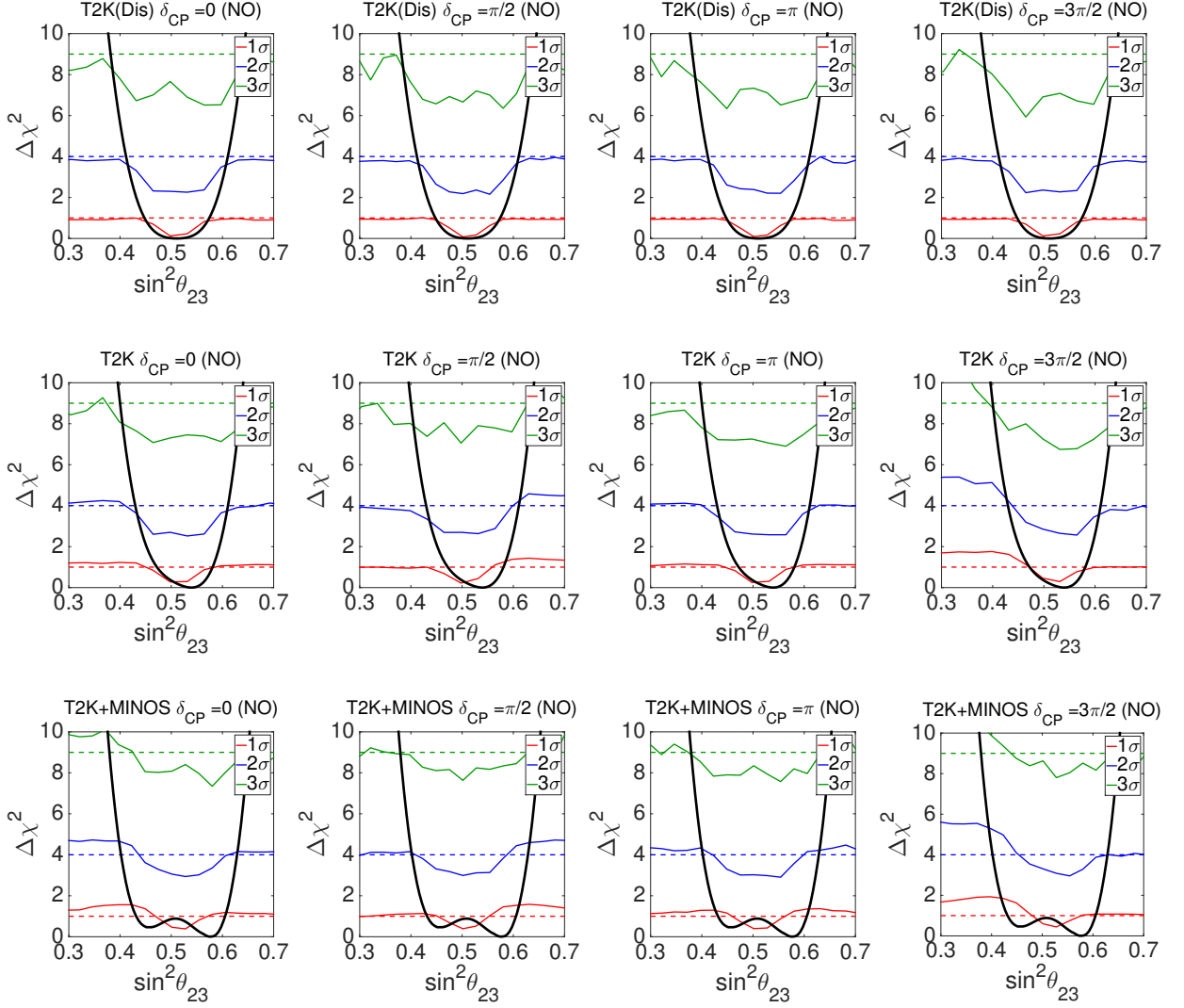
Figure 5: The cut levels $t_{\mathrm{MC}}(\mathrm{CL}, \Theta)$ for $\Delta\chi^2(\theta_{23})$ for 1σ (red), 2σ (blue), 3σ (green). Dashed lines indicate the Gaussian approximation t_{χ^2}. Top row: T2K disappearance only, middle row: T2K disappearance and appearance, bottom row: combined T2K and MINOS data. The columns correspond to $\delta_{\mathrm{CP}}^{\mathrm{true}} = 0, \pi/2, \pi, 3\pi/2$ from left to right, and we always assume a true normal mass ordering with $\Delta m^2_{32} = 2.4 \times 10^{-3}\,\mathrm{eV}^2$. The black solid curve shows $\Delta\chi^2(\theta_{23})$ using the observed data (same curve in the 4 panels in each row).

θ_{23}, and within a frequentist framework it is not possible to decide which range for θ_{23} has to be considered.[5] This problem is solved by considering 2-dimensional confidence regions in both parameters. Before presenting those in section 4.3 below, we proceed now by discussing the 1-dimensional intervals for θ_{23}.

4.2 One-dimensional intervals for θ_{23}

The cut levels for $\Delta\chi^2(\theta_{23})$ obtained from our MC simulation are shown in fig. 5 for T2K disappearance data only (top row), for T2K disappearance and appearance data (middle

---

[5]A similar discussion in the context of the mass ordering determination can be found in ref. [41].

row), and T2K and MINOS combined (bottom row). The columns of panels correspond to different values of the CP phase $\delta_{\rm CP}$ used to generate the pseudo data. We make the following observations:

1. For non-maximal values $|\sin^2\theta_{23} - 0.5| \gtrsim 0.05$, the test statistic $\Delta\chi^2(\theta_{23})$ is distributed approximately as a χ^2 with 1 dof, according to the Gaussian approximation.

2. For maximal values $\sin^2\theta_{23} \simeq 0.5$ the cut levels are somewhat reduced compared to the Gaussian case. This is the manifestation of the physical boundary $\sin^2 2\theta_{23} \leq 1$, which reduces the freedom provided by the parameter θ_{23}.

3. Disappearance data only (top row) is insensitive to the true value of the phase $\delta_{\rm CP}$.

4. When appearance data are included we observe a slight increase of the cut levels compared to the Gaussian limit for certain combinations of θ_{23} and $\delta_{\rm CP}^{\rm true}$, namely for $\delta_{\rm CP}^{\rm true} = \pi/2$, $\sin^2\theta_{23} \gtrsim 0.6$ and for $\delta_{\rm CP}^{\rm true} = 3\pi/2$, $\sin^2\theta_{23} \lesssim 0.4$. These are the same regions where we have noted also an increase in the cut levels for $\Delta\chi^2(\delta_{\rm CP})$ in the previous section, corresponding to the cases of strong octant degeneracy, see panels (d) and (f) in fig. 3.

5. The behavior of the cut levels is basically unchanged by adding MINOS data to T2K. However, the $\Delta\chi^2(\sin^2\theta_{23})$ from the observed data (black curves in the plots) is slightly disfavoring maximal mixing. While the confidence intervals for $\sin^2\theta_{23}$ based on the MC from T2K happen to be very similar to the χ^2 approximation, some deviations from the Gaussian limit occur when MINOS data are added, since the observed $\Delta\chi^2$ is pushed somewhat into regions where $t_{\rm MC}$ differs from t_{χ^2}.

Fig. 5 corresponds to assuming a normal ordering for generating the MC data. Very similar results are obtained also for inverted ordering. We have also investigated the distribution of the 1-dimensional $\Delta\chi^2$ for Δm_{23}^2 and we have found very good agreement with the Gaussian limit, independent of any other parameters for any combination of T2K and/or MINOS data. This reflects the very robust determination of Δm_{32}^2 by the ν_μ disappearance spectral data.

4.3 Two-dimensional confidence regions

As we have seen above, the distribution of the test statistic for $\delta_{\rm CP}$ depends significantly on the unknown true value of θ_{23}. Hence, treating θ_{23} as nuisance parameter leads to the unpleasant result that the confidence intervals for $\delta_{\rm CP}$ cannot be stated independently of the true value of θ_{23}. One way to deal with such a situation in a frequentist framework is to consider two-dimensional confidence regions of both parameters, keeping in mind that the interpretation of the results is different.

We consider the test statistic $\Delta\chi^2(\delta_{\rm CP}, \theta_{23})$ defined in eq. (5) and simulate the distribution of this statistic for a grid of true values in the $(\delta_{\rm CP}, \theta_{23})$ plane. Then in each point in this plane the observed value of $\Delta\chi^2(\delta_{\rm CP}, \theta_{23})$ can be compared to the distribution from the MC to decide whether this point is included in the confidence region at a given CL. The results of such an analysis are shown in fig. 6 for T2K only (left panel) and the T2K + MINOS combination (right panel). The thick solid curves indicate the confidence regions in the space of $\delta_{\rm CP}$ and $\sin^2\theta_{23}$ at 1σ, 2σ, and 3σ CL based on the MC simulation. They can be compared to the thin dotted curves, which indicate the regions obtained under the Gaussian

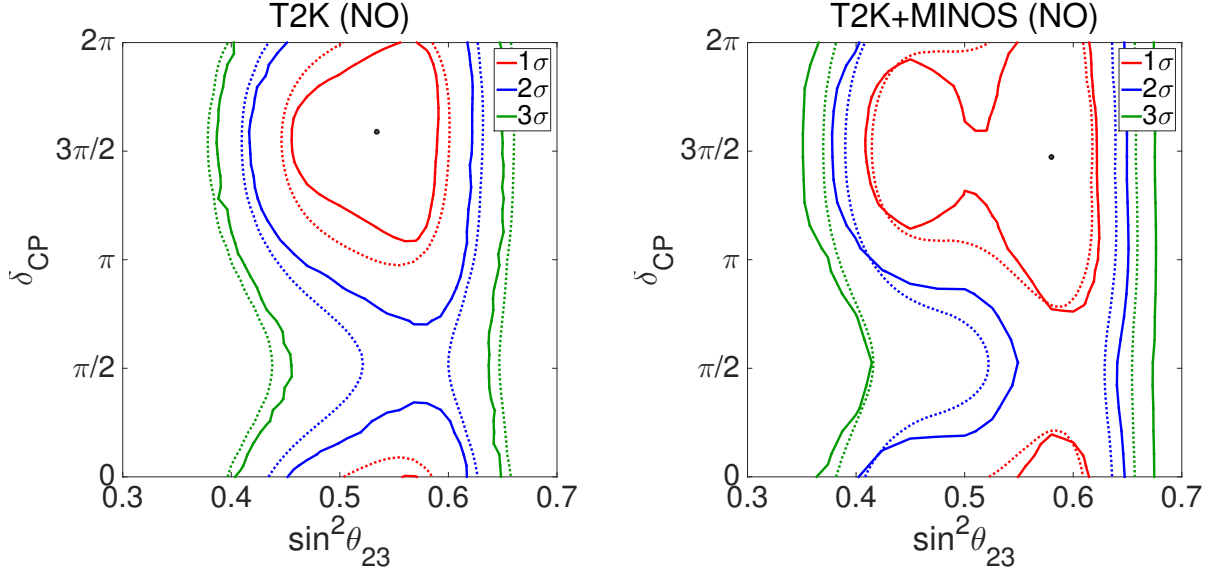Figure 6: Two-dimensional confidence regions at 1σ (red), 2σ (blue), 3σ (green) in the $(\sin^2\theta_{23}, \delta_{CP})$ plane for T2K (left panel) and T2K + MINOS (right panel). Solid curves correspond to the MC simulation, whereas dotted curves correspond to the Gaussian approximation. For the MC we assume a true normal mass ordering with $\Delta m_{32}^{2\,\text{true}} = 2.4 \times 10^{-3}\,\text{eV}^2$, while for the fit we minimize with respect to Δm_{32}^2 including its sign.

approximation, i.e. using the cut levels t_{χ^2} obtained from the χ^2 distribution for 2 dof. We observe that regions obtained from the MC are relatively close to the Gaussian limit. We conclude that the 2-dimensional test statistic has "better" statistical properties than the 1-dimensional one, where the θ_{23} dependence is profiled out.

This result is further illustrated in fig. 7, which shows sections through the 2-dimensional distribution. Those results should not be confused with the ones shown in figs. 1 and 2, where the χ^2 is minimised with respect to θ_{23}, whereas here we keep it fixed at the assumed true value (eq. (4) versus (5)). The MC curves in fig. 7 should be compared to the corresponding cut values t_{χ^2} for a χ^2 distribution with 2 dof, indicated by the dashed lines in the figure. We observe that the MC cut levels are close to the Gaussian limit. For $\sin^2\theta_{23} \simeq 0.5$ we observe somewhat smaller t_{MC} values compared to the Gaussian ones, due to the physical boundary $\sin^2 2\theta_{23} \leq 1$ (visible in the right panel in fig. 7). In all cases the variation of the MC cut levels with δ_{CP} as well as with θ_{23} is significantly reduced compared to the 1-dimensional case. We conclude that for present data, the Gaussian approximation to derive confidence regions is more reliable in the $(\delta_{CP}, \theta_{23})$ plane, whereas 1-dimensional confidence intervals for the CP phase δ_{CP} suffer from large deviations from Gaussianity.

For figs. 6 and 7 we have assumed a true normal mass ordering to generate the MC data. The corresponding plots for a true inverted ordering are very similar. We can use the 2-dimensional confidence regions also to quantify the rejection of $\delta_{CP} = \pi/2$ by looking for the largest CL for which the $(\delta_{CP}, \sin^2\theta_{23})$ confidence regions do not contain $\delta_{CP} = \pi/2$. In this way we find from the MC calculation that combined T2K and MINOS data allow to reject $\delta_{CP} = \pi/2$ at the 81.8% (83.9%) CL assuming a true normal (inverted) mass ordering. We
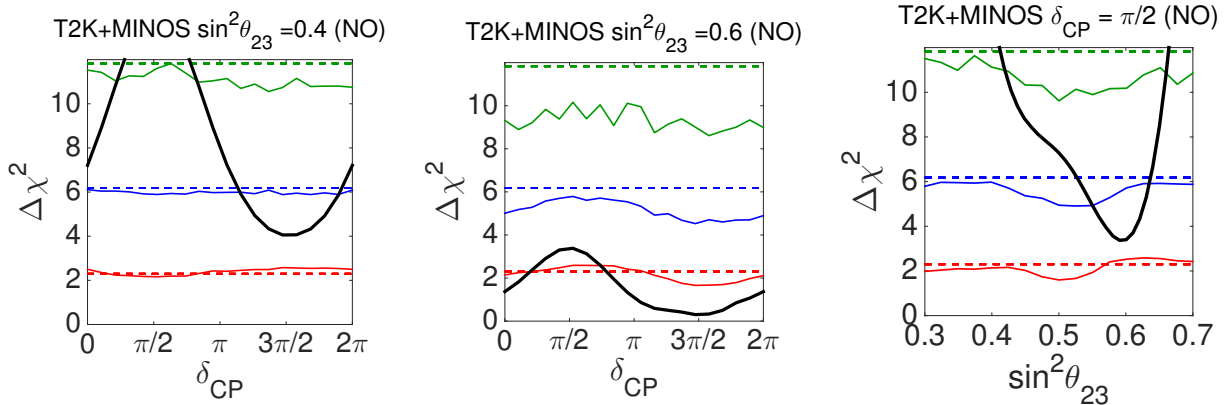
Figure 7: Sections through the two-dimensional $\Delta\chi^2(\delta_{CP}, \theta_{23})$ distribution for combined T2K and MINOS data at constant θ_{23} with $\sin^2\theta_{23} = 0.4$ (left panel) and $\sin^2\theta_{23} = 0.6$ (middle panel), and at constant $\delta_{CP} = \pi/2$ (right panel). Solid curves correspond to t_{MC} for 1σ (red), 2σ (blue), 3σ (green). Dashed lines indicate t_{χ^2} for a χ^2 distribution with 2 dof. The black solid curves show $\Delta\chi^2(\delta_{CP}, \theta_{23})$ using the observed data.

have $\Delta\chi^2(\delta_{CP} = \pi/2) = 3.37$, which in the Gaussian approximation for 2 dof corresponds to the 81.5% CL. We note that the MC results for NO and IO are very similar and close to the one obtained in the Gaussian limit. In contrast, the corresponding rejection confidence levels based on 1-dimensional confidence intervals for δ_{CP} vary strongly with the true mass ordering and true θ_{23} and differ significantly from the Gaussian limit (see discussion in section 4.1). Note that the interpretation of the rejection confidence levels based on 1-dimensional or 2-dimensional confidence regions is different.

In Fig. 8 we show the 2-dimensional Feldman-Cousins confidence regions in the plane of $\sin^2\theta_{23}$ and Δm_{32}^2 for the combined T2K and MINOS data. We observe that they agree quite well with the standard Gaussian approximation. For generating the MC data we have assumed here $\delta_{CP} = 3\pi/2$, but the results are very similar for other true values of δ_{CP}.

5 Increased exposure and T2K anti-neutrino data

Next we are going to discuss how this situation will change in the near future, for increased exposure in T2K or when data on anti-neutrinos become available. Following the T2K collaboration [42] we consider an exposure of 7.8×10^{21} protons-on-target (p.o.t.), approximately a factor 12 larger than the current exposure. We consider two cases, either using all of this exposure for neutrino data or equally sharing the exposure between neutrino and anti-neutrino running. We depart from the code used for our present-data T2K analysis (see section 2), and scale the spectrum normalization such that we can reproduce the expected number of events given in tables 4 and 5 in [42]. This is a rough approximation, especially for the anti-neutrino case, where we ignore the (rather substantial) contribution from the neutrino component in the beam. Despite those simplifications we can reproduce accurately the event spectra from fig. 2 in [42], as well as the sensitivity plots based on Asimov data given in ref. [42]. Note that total event numbers are still relatively small, 210 (260) events for
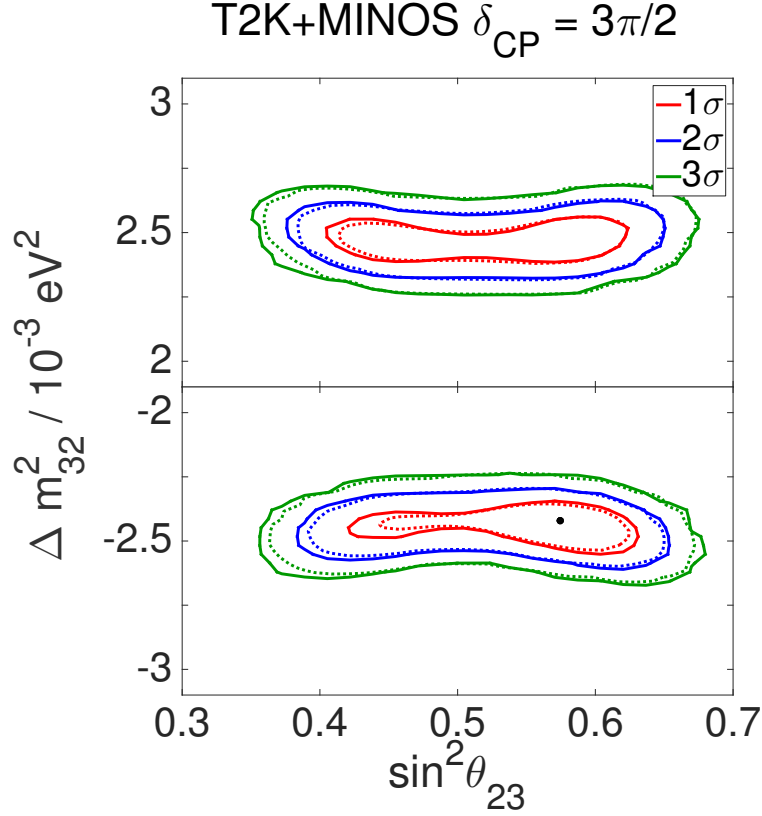
Figure 8: Two-dimensional confidence regions at 1σ (red), 2σ (blue), 3σ (green) in the $(\theta_{23}, \Delta m_{32}^2)$ plane for T2K + MINOS data. Solid curves correspond to the MC simulation, whereas dotted curves correspond to the Gaussian approximation. For the MC we assume a true value of $\delta_{CP} = 3\pi/2$, while for the fit we minimize with respect to δ_{CP}.

neutrinos and 49 (35) events for antineutrinos for $\delta_{CP} = 0$ $(3\pi/2)$, implying large statistical errors. Our approximate implementation suffices to study the expected statistical properties of the test statistics. A detailed and accurate sensitivity calculation is beyond the scope of this work.

In fig. 9 we show the cut levels for $\Delta\chi^2(\delta_{CP})$ for different assumptions about the true θ_{23} and mass ordering, both for neutrino-only data and for combining neutrino and anti-neutrino data. We find that in both cases a significant dependence on the unknown true values of θ_{23} and the mass ordering remains. In particular, the significance of excluding values of $\delta_{CP} \simeq \pi/2$ or $3\pi/2$ will vary quite strongly. The locations of the dips in the cut levels follow the pattern discussed in section 3. The regions of low cut levels for neutrino data-only visible for NO, $\sin^2\theta_{23} = 0.6$, $\delta_{CP} \simeq 3\pi/2$ (upper right panel) and IO, $\sin^2\theta_{23} = 0.4$, $\delta_{CP} \simeq \pi/2$ (lower left panel) correspond to the regions of maximal and minimal oscillation probability indicated in the first two lines of eq. (14). Also the dips in the middle panels of fig. 9 (for $\sin^2\theta_{23} = 0.5$) follow this argument.

We also observe from the figure that in many cases using anti-neutrino data has only a small impact on the distribution of the test statistics. The only exceptions are $\sin^2\theta_{23} = 0.4$, NO (upper left) and $\sin^2\theta_{23} = 0.6$, IO (lower right). In those cases cut levels are raised close to the Gaussian limit for $0 \leq \delta_{CP} \leq \pi$ or $\pi \leq \delta_{CP} \leq 2\pi$, respectively. In the opposite regions

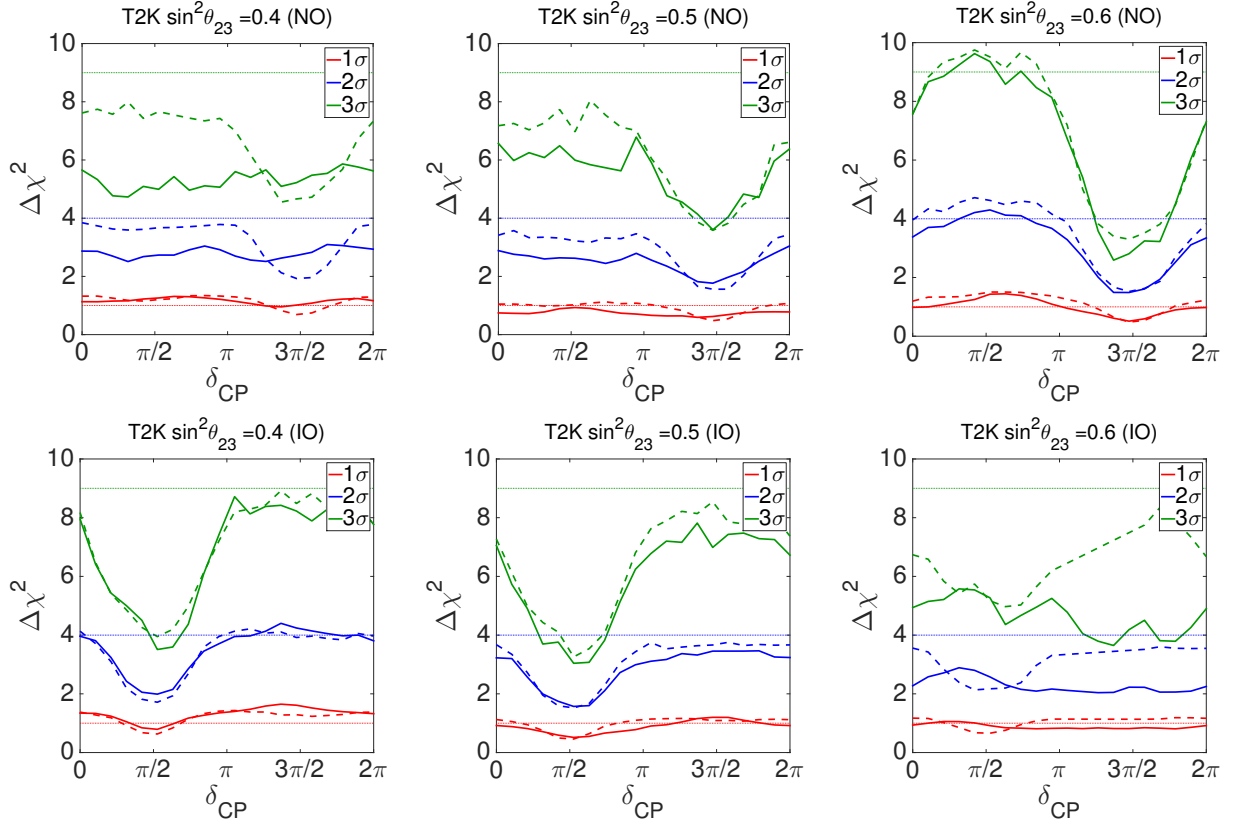Figure 9: The cut levels $t_{\mathrm{MC}}(\mathrm{CL}, \Theta)$ for $\Delta\chi^2(\delta_{\mathrm{CP}})$ for future T2K data of 7.8×10^{21} p.o.t. for 1σ (red), 2σ (blue), 3σ (green). Solid curves assume 100% neutrino running, dashed curves are for 50% neutrino and anti-neutrino data, each. Left, middle, right panels correspond to $\sin^2\theta_{23}^{\mathrm{true}} = 0.4, 0.5, 0.6$, respectively. We take $|\Delta m_{32}^2|^{\mathrm{true}} = 2.4 \times 10^{-3}\,\mathrm{eV}^2$ and for the upper (lower) row we have assumed a true normal (inverted) mass ordering. The mass ordering ambiguity is included in the minimisation.

of δ_{CP} for those cases cut levels remain low also when anti-neutrino data are added. For those regions, there is an octant degenerate solution which basically destroys any sensitivity to δ_{CP}, see panel (d) in fig. 3, while adding information from anti-neutrinos significantly helps in resolving this degeneracy which increases the sensitivity to δ_{CP}. Those regions correspond to the minimal and maximal values for the anti-neutrino oscillation probability indicated in the last two lines of eq. (14), explaining the good sensitivity of anti-neutrino data for those cases.

In fig. 10 we show the same analysis as in fig. 9, except that we assume that the mass ordering is known (determined independently by some other experiment). Hence, we do not minimize with respect to the mass ordering, but restrict the fit to the assumed true ordering. We observe that the sign(Δm_{32}^2) degeneracy has a large impact on the distribution of the test statistics and cut levels become much closer to the Gaussian approximation.[6] This is different to the situation we find for current data discussed in section 4, where the mass ordering degeneracy has only a small impact on the distribution of the test statistics. For

---

[6]This is an example where the presence of the degeneracy decreases the effective number of dof, contrary to the cases discussed previously, where the presence of a degeneracy increases the number of dof.
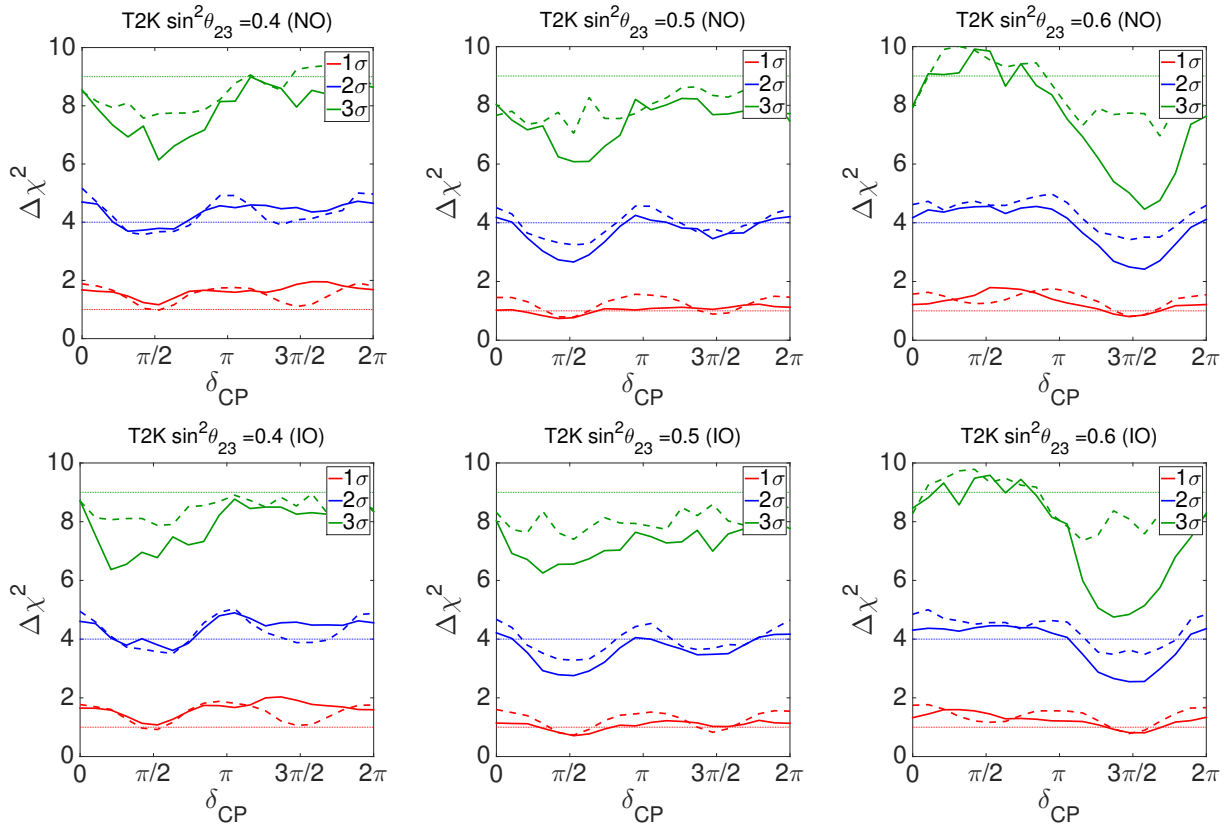
Figure 10: Same as figure 9, but assuming that the mass ordering is known, i.e. restricting the fit to the assumed true ordering.

neutrino-only data, there are still a few cases of large deviations from Gaussianity, whereas the combined neutrino and anti-neutrino data lead to cut levels close to the χ^2 limit (see dashed curves in fig. 10). We still observe small dips for $\delta_{CP} \simeq \pi/2$ and $3\pi/2$ (now pretty symmetric around $\delta_{CP} = \pi$), which can be traced back to the $\sin\delta_{CP}$ dependence of the probabilities, leading to a reduction of the dof for those values of δ_{CP}. Note also that in some cases we find now cut levels which are even higher than the Gaussian ones. A possible explanation for this behavior can be found in the discussion in section 3, see item 1.

6 Discussion and conclusions

We have studied in detail the information we can obtain on the leptonic CP phase δ_{CP} from current data, focusing on the robustness of frequentist confidence regions, by performing a Monte Carlo simulation of the data from the T2K and MINOS experiments. We attempt to quantify the current preference for $\delta_{CP} \simeq 3\pi/2$ over $\delta_{CP} \simeq \pi/2$. We have focused on the interplay of the main unknown parameters, namely δ_{CP}, θ_{23}, and the neutrino mass ordering. Our findings can be summarized as follows.

The distribution of the $\Delta\chi^2$ test statistic used for 1-dimensional confidence intervals for δ_{CP} shows large deviations from the Gaussian limit. In particular, it strongly depends on the unknown true value of θ_{23}. This introduces an ambiguity in the confidence intervals for δ_{CP},

see fig. 4. While the 1σ interval for δ_{CP} is relatively stable and close to the Gaussian approximation, at higher confidence level large variations occur. In particular, the CL with which values of $\delta_{\mathrm{CP}} \simeq \pi/2$ are disfavored ranges from 2σ to 3σ, depending on θ_{23} and the mass ordering. We can trace back the origin of those results to the complicated non-linear parameter dependence of the relevant oscillation probabilities (trigonometric dependence of δ_{CP} and θ_{23}-octant and mass ordering degeneracies), combined with the rather poor sensitivity of current data to δ_{CP}.

We conclude that one should not use the Gaussian approximation when making statements about δ_{CP} based on the 1-dimensional $\Delta\chi^2$ test statistic. Typically the "true" confidence levels obtained from the Monte Carlo simulation lead to more restrictive confidence intervals and to stronger rejections of values of δ_{CP} around $\pi/2$. In this sense the use of the Gaussian approximation is conservative. We have shown that the Gaussian approximation is better justified for 2-dimensional confidence regions in the plane of δ_{CP} and $\sin^2\theta_{23}$, see fig. 6. In particular, the dependence of the $\Delta\chi^2$ distribution on the true values of δ_{CP} and θ_{23} is much less severe in the 2-dimensional case. The 2-dimensional confidence region in the $(\delta_{\mathrm{CP}}, \theta_{23})$ plane for combined T2K and MINOS data does not include $\delta_{\mathrm{CP}} = \pi/2$ up to the 81.8% (83.9%) CL assuming a true normal (inverted) mass ordering.[7] Those values are close to the CL of 81.5% obtained under the Gaussian approximation.

We have considered also the 1-dimensional confidence intervals for θ_{23} and Δm^2_{32}. For θ_{23} we find approximately Gaussian behavior, with some deviations around maximal mixing, see fig. 5. This is a manifestation of the boundary $\sin^2 2\theta_{23} \leq 1$, which implies that the derivative of the event rates predicted for the disappearance channel with respect to θ_{23} is zero for $\theta_{23} = 45°$. The test statistic for Δm^2_{32} has a distribution very close to the Gaussian limit, as well as the 2-dimensional confidence regions in the $(\sin^2\theta_{23}, \Delta m^2_{32})$ plane.

In section 5 we have studied the distribution of the 1-dimensional $\Delta\chi^2$ test statistic for δ_{CP} assuming an increased exposure for T2K of 7.8×10^{21} protons-on-target, roughly a factor 12 larger than current exposure, where we consider also the possibility of using half of this exposure for anti-neutrino running. We find that even in this case large deviations from the Gaussian behavior can be expected. Typically reduced cut levels for the $\Delta\chi^2$ are obtained around either $\delta_{\mathrm{CP}} \simeq \pi/2$ or $3\pi/2$, depending on the unknown true value of θ_{23} and the mass ordering. Close to Gaussian behavior is only obtained for neutrino plus anti-neutrino running and assuming that the neutrino mass ordering is known.

Let us mention that in the global fit of all oscillation data also SuperKamiokande atmospheric neutrino data contribute to the determination of θ_{23} and to a small extent also of δ_{CP}, see ref. [17] for a discussion. Ideally a combined MC simulation of long-baseline and atmospheric neutrino data should be performed, which however, is not feasible due to the numerical complexity of the atmospheric neutrino fit. Since atmospheric neutrino data play only a subleading role for δ_{CP} we expect that the results presented here would not be modified substantially by including atmospheric neutrinos.

For the investigation of near term future data in section 5 we have not considered the NOvA experiment [43, 44], from which data will become available during the next years. In general we expect improved behaviour of the relevant test statistics, since complementary data from NOvA may help to resolve some of the degeneracies (see, e.g. refs. [28, 29] for recent

---

[7]When first data [34] from T2K on $\bar{\nu}_\mu \to \bar{\nu}_e$ are included, the corresponding numbers are 86.3% (89.2%) CL, see appendix.

studies), which—as we have shown—play a crucial role for the deviations from Gaussianity. The same is true also for experiments aiming to determine the neutrino mass ordering, see ref. [41] for a discussion and references. An exhaustive investigation of the expected statistical properties of future data is beyond the scope of this work. Some discussion along those lines in the context of CP violation can be found in ref. [32]. It will be an interesting topic for future work to study sensitivities to δ_{CP} of combined data from NOvA, T2K, and other upcoming experiments based on the true distributions of the relevant test statistics.

Acknowledgements

A Impact of first anti-neutrino data from T2K

After completion and submission of this paper, the first anti-neutrino results from T2K were presented at the EPS HEP 2015 conference [34]. In this appendix we show the impact of those data on the determination of δ_{CP}.

The results presented in [34] corresponds to about 4×10^{20} p.o.t. in the anti-neutrino mode with 3 observed events in the appearance channel.[8] The expected background (NC and other) is 1.17 events, while the predicted signal from $\overline{\nu}_\mu \to \overline{\nu}_e$ ($\nu_\mu \to \nu_e$) induced events ranges from 2 to 4 (0.3 to 0.6) events, depending on δ_{CP} and the mass ordering. We could reproduce the predicted number of events given in [34] with good accuracy, which allows us to include those data and combine it with the other data used in this work. We use a χ^2 as given in eq. (2) with just one bin (only the total number of events is fitted), taking into account the background expectation, as well as oscillated anti-neutrino and neutrino event predictions.

Clearly the statistical significance of those results is poor, since 3 observed events are even consistent with the background only hypothesis of about 1.17 events (no oscillation induced events at all) at slightly more than 1σ. Therefore we expect that those data will change $\Delta\chi^2$ by about 1 unit. Nevertheless, anti-neutrinos carry complementary information to the neutrino data and therefore it may be interesting to investigate the impact of those results.

Fig. 11 shows the impact of the anti-neutrino data on the 1-dimensional confidence interval for δ_{CP}. The solid curves, both for the cut levels as well as for $\Delta\chi^2(\delta_{\mathrm{CP}})$ are without T2K anti-neutrinos and are identical to fig. 2, while the dashed curves include the information from the T2K anti-neutrino events. Comparing the black solid and dashed curve we find that $\Delta\chi^2$ for $\delta_{\mathrm{CP}} = \pi/2$ is increased by about 0.75. However, also the cut levels are increased by a similar amount and hence the significance of rejecting $\delta_{\mathrm{CP}} = \pi/2$ is hardly affected. For $\sin^2\theta_{23} = 0.4$ the effect is most pronounced and in those cases the significance actually decreases, despite the increased $\Delta\chi^2$.

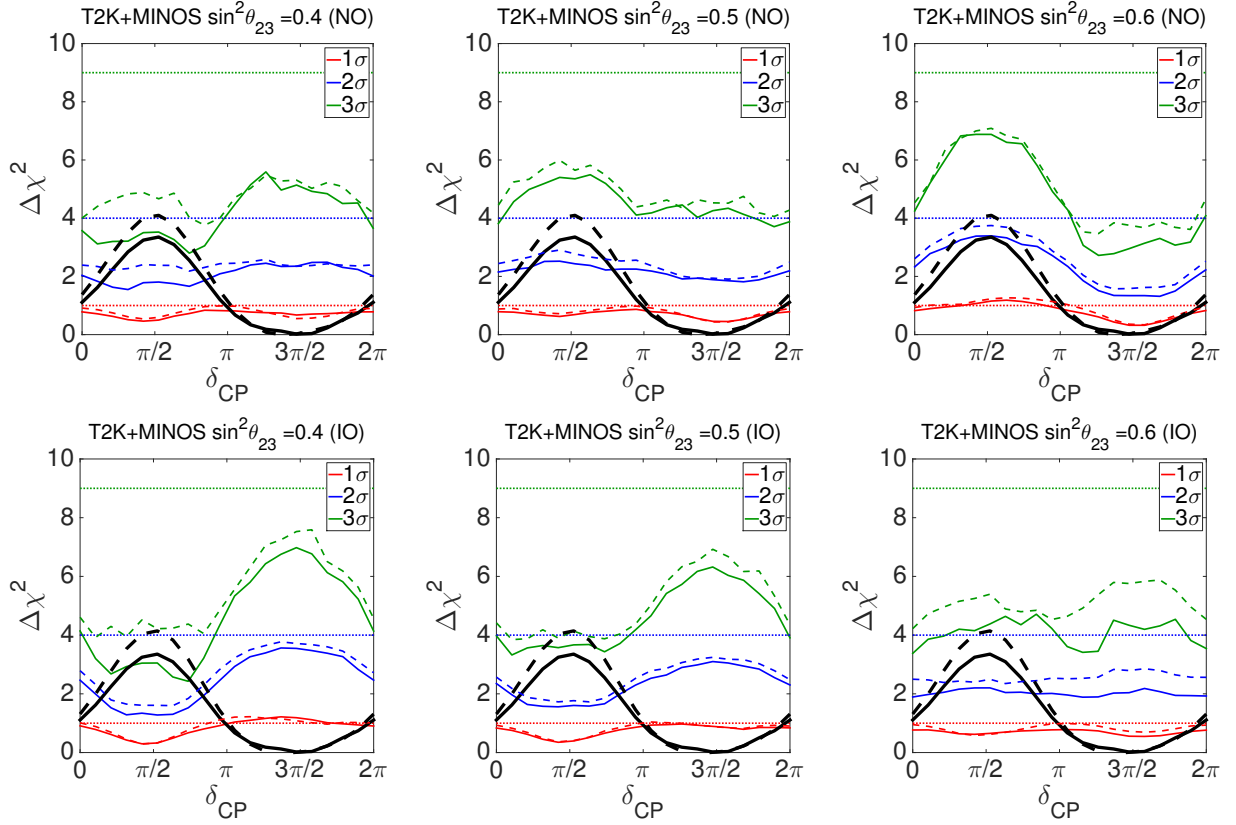[8]We do not consider the disappearance channel here.

Figure 11: The cut levels $t_{\mathrm{MC}}(\mathrm{CL}, \Theta)$ for $\Delta\chi^2(\delta_{\mathrm{CP}})$ from combined T2K and MINOS data for 1σ (red), 2σ (blue), 3σ (green). Left, middle, right panels correspond to $\sin^2\theta_{23}^{\mathrm{true}} = 0.4, 0.5, 0.6$, respectively. We take $|\Delta m_{32}^{2\ \mathrm{true}}| = 2.4 \times 10^{-3}\,\mathrm{eV}^2$ and for the upper (lower) row we have assumed a true normal (inverted) mass ordering. The black curves show $\Delta\chi^2(\delta_{\mathrm{CP}})$ using the observed data (same curves in all panels). Solid curves correspond to the data given in tab. 1 (no T2K anti-neutrinos, same as in fig. 2), dashed curves include T2K anti-neutrino data.

In fig. 12 we show the 2-dimensional confidence region in the $(\delta_{\mathrm{CP}}, \theta_{23})$ plane for T2K as well as T2K+MINOS data, including the T2K anti-neutrino events in both cases. This figure should be compared with fig. 6, which shows the corresponding regions without T2K anti-neutrinos. As expected the difference is small, with the size of the confidence region being slightly decreased due to the new data. Including T2K anti-neutrinos we find that combined T2K and MINOS data allow to reject $\delta_{\mathrm{CP}} = \pi/2$ at the 86.3% (89.2%) CL assuming a true normal (inverted) mass ordering, to be compared with 81.8% (83.9%) CL without T2K anti-neutrinos. For $\Delta\chi^2(\delta_{\mathrm{CP}} = \pi/2)$ we find now a value of 4.27, which in the Gaussian approximation for 2 dof corresponds to the 88.2% CL.

References

[1] Super-Kamiokande, Y. Fukuda *et al.*, *Evidence for oscillation of atmospheric neutrinos*, Phys. Rev. Lett. **81**, 1562 (1998), hep-ex/9807003.
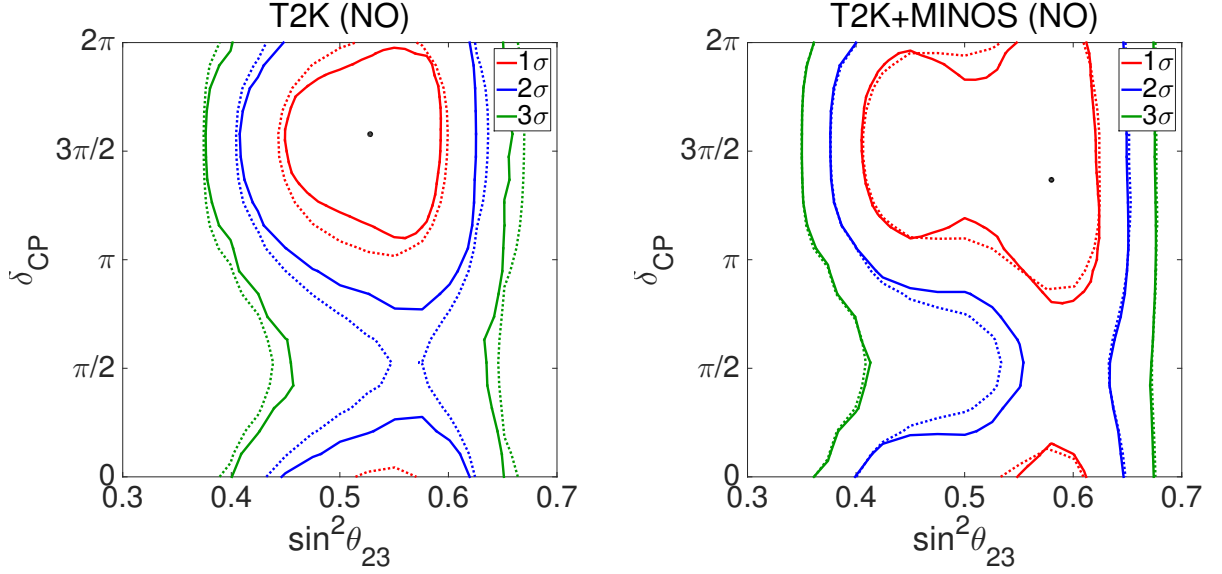
Figure 12: Two-dimensional confidence regions at 1σ (red), 2σ (blue), 3σ (green) in the $(\sin^2\theta_{23}, \delta_{\mathrm{CP}})$ plane for T2K (left panel) and T2K + MINOS (right panel), including T2K anti-neutrino data in both cases. Solid curves correspond to the MC simulation, whereas dotted curves correspond to the Gaussian approximation. For the MC we assume a true normal mass ordering with $\Delta m_{32}^{2\,\mathrm{true}} = 2.4 \times 10^{-3}\,\mathrm{eV}^2$, while for the fit we minimize with respect to Δm_{32}^2 including its sign.

[2] Super-Kamiokande, R. Wendell *et al.*, *Atmospheric neutrino oscillation analysis with sub-leading effects in Super-Kamiokande I, II, and III*, Phys.Rev. **D81**, 092004 (2010), 1002.3471.

[3] SNO, Q. R. Ahmad *et al.*, *Direct evidence for neutrino flavor transformation from neutral-current interactions in the Sudbury Neutrino Observatory*, Phys. Rev. Lett. **89**, 011301 (2002), nucl-ex/0204008.

[4] KamLAND, T. Araki *et al.*, *Measurement of neutrino oscillation with KamLAND: Evidence of spectral distortion*, Phys. Rev. Lett. **94**, 081801 (2005), hep-ex/0406035.

[5] MINOS, P. Adamson *et al.*, *Measurement of Neutrino Oscillations with the MINOS Detectors in the NuMI Beam*, Phys. Rev. Lett. **101**, 131802 (2008), 0806.2237.

[6] MINOS, P. Adamson *et al.*, *Electron neutrino and antineutrino appearance in the full MINOS data sample*, Phys.Rev.Lett. **110**, 171801 (2013), 1301.4581.

[7] MINOS, P. Adamson *et al.*, *Measurement of Neutrino and Antineutrino Oscillations Using Beam and Atmospheric Data in MINOS*, Phys.Rev.Lett. **110**, 251801 (2013), 1304.6335.

[8] Double Chooz, Y. Abe *et al.*, *Reactor electron antineutrino disappearance in the Double Chooz experiment*, Phys.Rev. **D86**, 052008 (2012), 1207.6632.

[9] Double Chooz, Y. Abe *et al.*, *Improved measurements of the neutrino mixing angle θ_{13} with the Double Chooz detector*, JHEP **1410**, 086 (2014), 1406.7763.

[10] RENO, J. Ahn *et al.*, *Observation of Reactor Electron Antineutrino Disappearance in the RENO Experiment*, Phys.Rev.Lett. **108**, 191802 (2012), 1204.0626.

[11] DAYA-BAY, F. An *et al.*, *Observation of electron-antineutrino disappearance at Daya Bay*, Phys.Rev.Lett. **108**, 171803 (2012), 1203.1669.

[12] Daya Bay, F. An *et al.*, *Spectral measurement of electron antineutrino oscillation amplitude and frequency at Daya Bay*, Phys.Rev.Lett. **112**, 061801 (2014), 1310.6732.

[13] T2K, K. Abe *et al.*, *Observation of Electron Neutrino Appearance in a Muon Neutrino Beam*, Phys.Rev.Lett. **112**, 061802 (2014), 1311.4750.

[14] T2K, K. Abe *et al.*, *Precise Measurement of the Neutrino Mixing Parameter θ_{23} from Muon Neutrino Disappearance in an Off-Axis Beam*, Phys.Rev.Lett. **112**, 181801 (2014), 1403.1532.

[15] T2K, K. Abe *et al.*, *Measurements of neutrino oscillation in appearance and disappearance channels by the T2K experiment with 6.6×10^{20} protons on target*, Phys.Rev. **D91**, 072010 (2015), 1502.01550.

[16] Particle Data Group, K. Olive *et al.*, *Review of Particle Physics*, Chin.Phys. **C38**, 090001 (2014).

[17] M. Gonzalez-Garcia, M. Maltoni, and T. Schwetz, *Updated fit to three neutrino mixing: status of leptonic CP violation*, JHEP **1411**, 052 (2014), 1409.5439.

[18] F. Capozzi *et al.*, *Status of three-neutrino oscillation parameters, circa 2013*, Phys.Rev. **D89**, 093018 (2014), 1312.2878.

[19] D. Forero, M. Tortola, and J. Valle, *Neutrino oscillations refitted*, Phys.Rev. **D90**, 093006 (2014), 1405.7540.

[20] N. Cabibbo, *Time Reversal Violation in Neutrino Oscillation*, Phys.Lett. **B72**, 333 (1978).

[21] S. M. Bilenky, J. Hosek, and S. Petcov, *On Oscillations of Neutrinos with Dirac and Majorana Masses*, Phys.Lett. **B94**, 495 (1980).

[22] V. D. Barger, K. Whisnant, and R. Phillips, *CP Violation in Three Neutrino Oscillations*, Phys.Rev.Lett. **45**, 2084 (1980).

[23] H. Nunokawa, S. J. Parke, and J. W. Valle, *CP Violation and Neutrino Oscillations*, Prog.Part.Nucl.Phys. **60**, 338 (2008), 0710.0554.

[24] G. Branco, R. G. Felipe, and F. Joaquim, *Leptonic CP Violation*, Rev.Mod.Phys. **84**, 515 (2012), 1111.5332.

[25] P. Huber, M. Lindner, T. Schwetz, and W. Winter, *First hint for CP violation in neutrino oscillations from upcoming superbeam and reactor experiments*, JHEP **11**, 044 (2009), 0907.1896.

[26] S. Prakash, S. K. Raut, and S. U. Sankar, *Getting the Best Out of T2K and NOvA*, Phys.Rev. **D86**, 033012 (2012), 1201.6485.

[27] S. K. Agarwalla, S. Prakash, S. K. Raut, and S. U. Sankar, *Potential of optimized NOvA for large θ_{13} & combined performance with a LArTPC & T2K*, JHEP **1212**, 075 (2012), 1208.3644.

[28] P. Machado, H. Minakata, H. Nunokawa, and R. Zukanovich Funchal, *What can we learn about the lepton CP phase in the next 10 years?*, JHEP **1405**, 109 (2014), 1307.3248.

[29] M. Ghosh, P. Ghoshal, S. Goswami, N. Nath, and S. K. Raut, *Identifying and resolving the degeneracies in neutrino oscillation parameters in current experiments*, (2015), 1504.06283.

[30] S. Wilks, *The large-sample distribution of the likelihood ratio for testing composite hypotheses*, Annals Math. Statist. **9**, no. 1, 60 (1938).

[31] T. Schwetz, *What is the probability that theta(13) and CP violation will be discovered in future neutrino oscillation experiments?*, Phys.Lett. **B648**, 54 (2007), hep-ph/0612223.

[32] M. Blennow, P. Coloma, and E. Fernandez-Martinez, *Reassessing the sensitivity to leptonic CP violation*, JHEP **1503**, 005 (2015), 1407.3274.

[33] G. J. Feldman and R. D. Cousins, *A Unified approach to the classical statistical analysis of small signals*, Phys.Rev. **D57**, 3873 (1998), physics/9711021.

[34] M. Ravonel, *Antineutrino oscillations with T2K*, (2015), talk at EPS HEP 2015, 22–29 July 2015, Vienna, Austria.

[35] A. Cervera *et al.*, *Golden measurements at a neutrino factory*, Nucl.Phys. **B579**, 17 (2000), hep-ph/0002108.

[36] M. Freund, *Analytic approximations for three neutrino oscillation parameters and probabilities in matter*, Phys.Rev. **D64**, 053003 (2001), hep-ph/0103300.

[37] G. L. Fogli and E. Lisi, *Tests of three flavor mixing in long baseline neutrino oscillation experiments*, Phys.Rev. **D54**, 3667 (1996), hep-ph/9604415.

[38] H. Minakata and H. Nunokawa, *Exploring neutrino mixing with low energy superbeams*, JHEP **10**, 001 (2001), hep-ph/0108085.

[39] H. Minakata and S. J. Parke, *Correlated, precision measurements of θ_{23} and δ using only the electron neutrino appearance experiments*, Phys.Rev. **D87**, 113005 (2013), 1303.6178.

[40] P. Coloma, H. Minakata, and S. J. Parke, *Interplay between appearance and disappearance channels for precision measurements of θ_{23} and δ*, Phys.Rev. **D90**, 093003 (2014), 1406.2551.

[41] M. Blennow, P. Coloma, P. Huber, and T. Schwetz, *Quantifying the sensitivity of oscillation experiments to the neutrino mass ordering*, JHEP **1403**, 028 (2014), 1311.1822.

[42] T2K, K. Abe *et al.*, *Neutrino oscillation physics potential of the T2K experiment*, PTEP **2015**, 043C01 (2015), 1409.7469.

[43] NOvA Collaboration, D. Ayres *et al.*, *NOvA: Proposal to build a 30 kiloton off-axis detector to study $\nu_\mu \to \nu_e$ oscillations in the NuMI beamline*, (2004), hep-ex/0503053.

[44] NOvA, R. Patterson, *The NOvA Experiment: Status and Outlook*, Nucl.Phys.Proc.Suppl. **235-236**, 151 (2013), 1209.0716.